PROPOSED SMPTE STANDARD

for Television —

MPEG-2 Video Elementary Stream Editing Information

1 Scope

This standard defines the MPEG video elementary stream (ES) information to facilitate seamless edits under defined circumstances. The video ES, as defined by the MPEG standards, are supplemented with additional information for professional studio applications. Supplementary information will be carried within the sequence header and the user data area of the video ES. This standard defines the data to be carried and the location of the data.

2 Normative references

The following standards contain provisions which, through reference in this text, constitute provisions of this standard. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this standard are encouraged to investigate the possibility of applying the most recent edition of the standards indicated below.

ANSI/SMPTE 293M-1996, Television — 720 × 483 Active Line at 59.94-Hz Progressive Scan Production — Digital Representation

ANSI/SMPTE 296M-1997, Television — 1280 × 720 Scanning, Analog and Digital Representation and Analog Interface

SMPTE 12M-1999, Television, Audio and Film — Time and Control Code

SMPTE 274M-1998, Television — 1920 × 1080 Scanning and Analog and Parallel Digital Interfaces for Multiple Picture Rates

SMPTE 309M-1999, Television — Transmission of Date and Time Zone Information in Binary Groups of Time and Control Code

SMPTE 326M, Television — SDTI Content Package Format (SDTI-CP)

SMPTE 327M, Television — MPEG-2 Video Recoding Data Set

SMPTE RP 186-1995, Video Index Information Coding for 525- and 625-Line Television Systems

SMPTE RP 202, Video Alignment for MPEG-2 Coding

ISO/IEC 13818-2:1996, Information Technology — Generic Coding of Moving Pictures and Associated Audio Information: Video

ITU-R BT.601-5 (10/95), Studio Encoding Parameters of Digital Television for Standard 4:3 and Widescreen 16:9 Aspect Ratios

ITU-R BT.656-4 (02/98), Interfaces for Digital Component Video Signals in 525-Line and 625-Line Television Systems Operating at the 4:2:2 Level of Recommendation ITU-R BT.601 [Part A]

3 MPEG ES

ISO/IEC 13818-2 defines the characteristics of the MPEG elementary stream (ES). When used in a professional environment where seamless edits (splices) are required, supplementary data needs to be carried by the bit stream.

3.1 General

Seamless frame accurate editing of compressed video is most easily accomplished with use of short GOP structures. Longer GOP structures can be edited by decoding and reencoding, by transcoding to shorter GOP structures, or with more involved processing, edited directly. The best approach will be determined by a range of application-specific considerations.

3.2 Sequence header

ISO/IEC 13818-2 does not define the repetition frequency of the sequence header. To be compliant with this standard, the sequence header shall exist at every I frame.

3.3 VBV_delay

An accurate vbv_delay value shall be carried in the video elementary stream. The value of 0xffff shall not be allowed. The vbv_delay value may be relied upon for remultiplexiing elementary streams into transport streams.

4 MPEG ES-syntax elements

The following syntax elements, and the functional descriptions, shall be inserted in the MPEG ES in the user data area:

4.1 V/H coding phase

The current implementation of MPEG does not specify the horizontal and vertical coding phase. This standard requires that the vertical and horizontal coding phase be known in order for decoding and peripheral equipment to process correctly the signal.

V and H coding information shall be included only for SDTV signals where the coding phase is not compliant with SMPTE RP 202. For HDTV signals, H/V coding phase information is defined by SMPTE RP 202.

4.2 Time code

Provision is made for the insertion of two time codes complying with SMPTE 12M. At least one time code, reference date time stamp (as defined in SMPTE 326M) shall be carried as a means of maintaining synchronization with other content or metadata streams. Carriage of a second time code is optional. Compliant decoders shall have the capability to decode the two time codes.

4.3 Picture order

Picture order information specifies the picture duration and is the equivalent to the PTS/DTS present in the MPEG transport stream. Picture order value is counted by field units. Picture order information will be used by multiplexers and remultiplexers to know the picture duration and to be able to parse the repeat_first_field information. In some cases, the latency of the system will be minimized using the picture order information.

4.4 Video index

Video index, as defined by SMPTE RP 186, shall be carried if present on the baseband signal. Information carried by video index should be preserved during any coding, recoding, editing, or transcoding process. It is envisioned that the data described in the forthcoming SMPTE metadata dictionary (SMPTE 335M) will be handled by the transport mechanism described in SMPTE 326M. These parametric data will include all of the parameters currently coded in video index, although the data representation of some items may be different.

4.5 Ancillary data

Data that is carried in the vertical interval of the baseband signal should be preserved. Ancillary data may consist of more than 23 consecutive zeros. To prevent this condition, a marker shall be inserted every 22 bits.

4.6 History data

History data, consisting of original and subsequent encoding parameters that may be useful in transcoding or reencoding, may be carried by the bit stream. SMPTE 327M defines the content of the history data information. History data may consist of more than 23 consecutive zeros. To prevent this condition, a marker shall be inserted every 22 bits.

4.7 User data

User data is defined by ISO/IEC 13818-2.

5 MPEG ES-syntax

MPEG_ES_editing_information is inserted in the sequence or/and picture user data area of the video ES. All data fields follow the unique 16-bit header which specifies this stream as being compliant with this standard. The syntax shall be flexible and extensible. All syntax elements are distinguished by data ID. Data ID specifies the length and characteristics of the following syntax element.

The V/H coding phase and defined control flags shall be transmitted in user_data of the sequence layer. Others shall be transmitted in user_data of the picture layer. In the picture layer, time_code() and picture_order() will be placed first because these are relatively small amounts of data and could be handled by software. The following data types, video_index(), ancillary_data(), and history_data() are positioned later in the bit stream because these data types potentially can be large in size and may require hardware support. User_data() shall be placed at the end (tail).

NOTE – All syntactic elements described in this standard follow the pseudo-code form, as defined in ISO/IEC 13818.

5.1 Data_ID

8 bits: This syntax element specifies the length and characteristics of the following; 0x00 is forbidden.

Data_ID	Data_type
00	FORBIDDEN
01	V-phase
02	H-phase
03	Time code 1
04	Time code 2
05	Picture order
06	Video index
07	Ancillary data
08	History data
...	...
80	Control flags
...	...
FF	User data

5.2 Higher syntactic structure

Syntax	Bits	Mnemonic
MPEG_ES_editing_information(i){		
SMPTE_header	16	"0x0001"
If (i == 0){/* follows sequence user data */		
While (nextbits() != "0x00"){		
If (nextbits() == "V-phase")		
V-phase()	...	
Else if (nextbits() == "H-phase")		
H-phase()	...	
Else if (nextbits() == "Control flags")		
Control flags()	...	
Else		
Break; / * Do nothing */		
}		
}		
Else if (i == 2){ /* follows picture user data */		
While (nextbits() != "0x00"){		
If (nextbits() == "Time code 1" || Nextbits() == "Time code 2")		
Time_code()	...	
Else if (nextbits() == "Picture_order"		
Picture_order()	...	
Else		
Break; /* Do nothing */		
}		
While (nextbits() != "0x00"){		
If (nextbits() == "Video index")		
Video_index()	...	
Else if (nextbits() == "Ancillary data")		
Ancillary_data()	...	
Else if (nextbits() == "History data")		
History_data()	...	
Else		
Break; / * Do nothing * /		
}		
}		
If (nextbits() == "User data")		
User_data()	...	
Next_start_code()		
}		

5.3 V coding phase

Syntax	Bits	Mnemonic
V-phase(){		
Data_ID	8	bslbf
V-phase	16	uimsbf
}		

V-phase (16 bits): V-phase is an unsigned integer that specifies the top line of the coded frame. When present, V-phase shall be carried in the user data of the sequence layer. For SDTV signals, V-phase shall be carried if the vertical coding phase is not compliant with SMPTE RP 202. For HDTV signals, V coding phase is defined by the SMPTE standard and the V-phase user data shall not be carried.

5.4 H coding phase

Syntax	Bits	Mnemonic
H-phase(){		
Data_ID	8	bslbf
H-phase	8	uimsbf
}		

H-phase (8 bits): H-phase is an unsigned integer that specifies the sample number of the first sample in a video line as defined in ITU-R BT.601. When present, H-phase shall be carried in the user data of the sequence layer. For SDTV signals, H-phase shall be carried if the horizontal coding phase is not compliant with SMPTE RP 202. For HDTV signals, H coding phase is defined by the SMPTE standard and the H-phase user data shall not be carried.

5.5 Control flags

Syntax	Bits	Mnemonic
Control flags(){		
Data_ID	8	bslbf
Picture_order_presence	1	uimsbf
Reserved	7	
}		

Picture_order_presence (1 bit): If this flag is set to 1, all picture order information shall be carried in the user data area of every picture. In this case, the multiplexer may use the picture order information for multiplexing, and low-delay multiplexing will be possible.

5.6 Time coding data

Syntax	Bits	Mnemonic
Time_code(){		
Data_ID	8	bslbf
Time_code [63..48]	16	uimsbf
Marker_bit	1	bslbf
Time_code [47..32]	16	uimsbf
Marker_bit	1	bslbf
Time_code [31..16]	16	uimsbf
Marker_bit	1	bslbf
Time_code [15..0]	16	uimsbf
Marker_bit	1	bslbf
Reserved_bits	4	bslbf
}		

Time_code (64 bits): The time-code format defined below shall comply with ANSI/SMPTE 12M. Two time codes may be carried. At least one time code (reference time/date stamp [SMPTE 309M]) shall be carried. All time codes specified shall be presented at the time the corresponding picture is established during the encoding process.

NOTE — The 64-bit contents of the time code are mapped as follows:

Syntax	Bits	Mnemonic
Time_code [63..0]{		
Color frame flag	1	bslbf
Drop frame flag (NTSC)/unused (PAL)	1	bslbf
TV frame tens	2	uimsbf
TV frame units	4	uimsbf
Field phase (NTSC)/binary group flag 0 (PAL)	1	bslbf
TV seconds tens	3	uimsbf
TV seconds units	4	uimsbf
Binary group flag 0 (NTSC)/binary group flag 2 (PAL)	1	bslbf
TV minutes tens	3	uimsbf
TV minutes units	4	uimsbf
Binary group flag 2 (NTSC)/field phase (PAL)	1	bslbf
Binary group flag 1 (NTSC)/binary group flag 1 (PAL)	1	bslbf
TV hours tens	2	uimsbf
TV hours units	4	uimsbf
2nd binary group	4	uimsbf
1st binary group	4	uimsbf
4th binary group	4	uimsbf
3rd binary group	4	uimsbf
6th binary group	4	uimsbf
5th binary group	4	uimsbf
8th binary group	4	uimsbf
7th binary group	4	uimsbf
}	4	uimsbf

5.7 Picture order

Syntax	Bits	Mnemonic
Picture_order(){		
Data_ID	8	bslbf
DTS_presence	1	bslbf
PTS_counter	7	uimsbf
If (DTS_presence == "1"){		
Marker_bits	1	bslbf
DTS_counter	7	uimsbf
}		
}		

DTS_presence (1 bit): If DTS_presence is set to 1, DTS_counter field shall exist.

PTS_counter (7 bits): This is a 7-bit unsigned integer which is equivalent to PTS counted by field units. This is a modulo 128 counter and shall increment according to picture duration taking into account repeat_first_field and reordering delay caused by B-picture.

DTS_counter (7 bits): This is a 7-bit unsigned integer which is equivalent to DTS counted by field units. This is a modulo 128 counter and shall increment according to picture duration taking into account repeat_first_field and reordering delay caused by B-picture.

Annex A presents two examples of PTS/DTS counter operation for long GOP and short GOP bit stream formats.

5.8 Video index

Syntax	Bits	Mnemonic
Video_index(){		
Data_ID	8	bslbf
Field_ID	2	bslbf
Line_number	14	uimsbf
Video_index_length	8	uimsbf
Marker_bits	1	bslbf
For (j=0; j<video_index_length; j++){		
Video_index_payload	22	uimsbf
Marker_bits	1	bslbf
}		
While (!bytealigned())		
Zero_bit	1	"0"
}		

Field_ID (2 bits): If the progressive_sequence flag is set to 0 (interlace), field_ID shall specify the field index counted by field unit in the presentation order of this picture. The counter of the earliest field shall be set to 0. If repeat_first_field is set to 0, the picture shall include two fields indexed as 0 and 1. If repeat_first_field is set to 1, the picture shall include three fields indexed as 0, 1, and 2.

If progressive_sequence flag is set to 1 (progressive), field_ID shall specify the frame index counted by progressive frame unit in the presentation order of this picture. The counter of the earliest frame shall be set to 0. If repeat_first_field and top_field_first are set to 0, the picture shall include only one progressive frame indexed as 0. If repeat_first_field is set to 1 and top_field_first are set to 0, the picture shall include two progressive frames indexed as 0 and 1. If repeat_first_field is set to 1 and top_field_first is set to 1, the picture shall include three progressive frames indexed as 0, 1, and 2.

Line_number (14 bits): This specifies the absolute frame-based line number of the video_index data; 0 is not allowed. The line number shall be specified according to the relevant video standard (ITU-R BT.656, SMPTE 274M, ANSI/SMPTE 293M, ANSI/SMPTE 296M).

Video_index_length (8 bits): This specifies the loop count of the following video_index_payload.

Video_index_payload (22 bits): These contain the payload of the video_index data. Video_index_payload is defined in SMPTE RP 186 for SDTV.

5.9 Ancillary data

Syntax	Bits	Mnemonic
Ancillary_data(){		
Data_ID	8	bslbf
Field_ID	2	bslbf
Line_number	14	uimsbf
Ancillary_data_length	16	uimsbf
Marker_bits	1	bslbf
For (j=0; j<ancillary_data_length; j++){		
Ancillary_data_payload	22	uimsbf
Marker_bits	1	bslbf
}		
While (!bytealigned())		
Zero_bit	1	"0"
}		

Field_ID (2 bits): If the progressive_sequence flag is set to 0 (interlace), field_ID shall specify the field index counted by the field unit in the presentation order of

this picture. The counter of the earliest field shall be set to 0. If repeat_first_field is set to 0, the picture shall include two fields indexed as 0 and 1. If repeat_first_field is set to 1, the picture shall include three fields indexed as 0, 1, and 2.

If the progressive_sequence flag is set to 1 (progressive), field _ID shall specify the frame index counted by the progressive frame unit in the presenttion order of this picture, The counter of the earliest frame shall be set to 0. If repreat_first_field and top_field_first are set to 0, the picture shall include only one progressive frame indexed as 0. If repeat_first_field is set to 1 and top_field_first is set to 0, the picture shall include two progressive frames indexed as 0 and 1. If repeat_first_field is set to 1 and top_field_first is set to 1, the picture shall include three progressive frames indexed as 0, 1, and 2.

Line_number (14 bits): This specifies the absolute frame-based line number of the ancillary data; 0 is not allowed. The line number shall be specified according to the relevant video standard (ITU-R BT.656, SMPTE 274M, ANSI/SMPTE 293M, ANSI/SMPTE 296M).

Ancillary_data_length (16 bits): This specifies the loop count of the following ancillary data_payload.

Ancillary_data_payload (22 bits): These contain the payload of the ancillary data.

5.10 History data

Syntax	Bits	Mnemonic
History_data(){		
Data_ID	8	bslbf
Marker_bits	1	bslbf
History_data_length	23	uimsbf
Marker_bits	1	bslbf
For (j=0; j<History_data_length; j++){		
History_data_payload	22	uimsbf
Marker_bits	1	bslbf
}		
While (!bytealigned())		
Zero_bit	1	"0"
}		

History_data_length (23 bits): This specifies the loop count of the following ancillary data_payload.

History_data__payload (22 bits): These bits contain the payload of the ancillary data. History data will be encapsulated in this field. The contents are defined in SMPTE 327M.

5.11 User data

Syntax	Bits	Mnemonic
User_data(){		
Data_ID	8	bslbf
While (nextbits() != "0x000001")		
User_data	8	uimsbf
}		
}		

User data (8 bits): The user data byte count is variable. User data shall not emulate the MPEG start codes (see ISO/IEC 13818-2). The end of user data is detected when the next start code appears. User data shall be placed at the end (tail) of the edit information.

Annex A (informative)
Usage of PTS and DTS_counter

A.1 Example 1: Long GOP

The following is an example of PTS and DTS_counter taken from the beginning of a video sequence. In this example, there are two coded B-frames between successive coded P-frames and also two coded B-frames between successive coded I- and P-frames and all pictures are frame pictures with a 3-2 pulldown operation.

At the encoder input:

Frame No.	1	2	3	4	5	6	7	8	9	10	11	12	13
Picture type	I	B	B	P	B	B	P	B	B	I	B	B	P
Repeat_first_field	1	0	1	0	1	0	1	0	1	0	1	0	1
Top_field_first	1	0	0	1	1	0	0	1	1	0	0	1	1

At the encoder output, in the coded bit stream, and at the decoder input:

Frame No.	1	4	2	3	7	5	6	10	8	9	13	11	12
Picture type	I	P	B	B	P	B	B	I	B	B	P	B	B
Repeat_first_field	1	0	0	1	1	1	0	0	0	1	1	1	0
Top_field_first	1	1	0	0	0	1	0	0	1	1	1	0	1
DTS_counter	125	0	3	5	8	10	13	15	18	20	23	25	28

At the decoder output:

Frame No.	1	2	3	4	5	6	7	8	9	10	11	12	13
Picture type	I	B	B	P	B	B	P	B	B	I	B	B	P
Repeat_first_field	1	0	1	0	1	0	1	0	1	0	1	0	1
Top_field_first	1	0	0	1	1	0	0	1	1	0	0	1	1
PTS_counter	0	3	5	8	10	13	15	18	20	23	25	28	30

A.2 Example 2: Short GOP

The following is an example of PTS and DTS_counter taken from the beginning of a short GOP video sequence. In this example, the counters are wrapping around at 20 instead of 128.

At the encoder input:

Frame No.	1	2	3	4	5	6	7	8	9
Picture type	I	B	I	B	I	B	I	B	I
Repeat_first_field	0	1	0	1	0	1	0	1	0
Top_field_first	1	1	0	0	1	1	0	0	1

At the encoder output, in the coded bit stream, and at the decoder input:

Frame No.	1	3	2	5	4	7	6	9	8
Picture type	I	I	B	I	B	I	B	I	B
Repeat_first_field	0	0	1	0	1	0	1	0	1
Top_field_first	1	0	1	1	0	0	1	1	0
DTS_counter	18	0	2	5	7	10	12	15	17

At the decoder output:

Frame No.	1	2	3	4	5	6	7	8	9
Picture type	I	B	I	B	I	B	I	B	I
Repeat_first_field	0	1	0	1	0	1	0	1	0
Top_field_first	1	1	0	0	1	1	0	0	1
PTS_counter	0	2	5	7	10	12	15	17	0

Annex B (informative)
Bibliography

SMPTE 291M-1998, Television — Ancillary Data Packet and Space Formatting

SMPTE 308M-1998, Television — MPEG-2 4:2:2 Profile at High Level

SMPTE 329M, Television — MPEG-2 Video Recoding Data Set — Compressed Stream Format

Forthcoming SMPTE 335M, Television — Metadata Dictionary

ISO/IEC 13818-1:1996, Information Technology — Generic Coding of Moving Pictures and Associated Audio Information: Systems